

לפעמים, אירוע אחד קטן חושף סדק גדול במערכת שלמה. Deloitte, אחת מחברות הייעוץ הגדולות בעולם, סיפקה לממשלת אוסטרליה דוח בעלות של 440 אלף דולר – דוח שנועד לבחון תהליכי בקרה פנים-ממשלתיים. אבל כשהמסמך הגיע לשולחנם של אנשי האקדמיה, משהו לא הסתדר. היו בו ציטוטים שלא קיימים, מקורות מומצאים, ואפילו פסק דין שמעולם לא נכתב. החוט המקשר בין כל הטעויות האלה? מודל הבינה המלאכותית, GPT-4o ששולב בהכנת הדוח מבלי שאיש טרח לבדוק את תוצריו לעומק. דלוייט תיאלץ כעת להשיב חלק מהכסף, אבל הכסף הוא רק הסיפור הקטן. הסיפור הגדול באמת הוא השאלה שמרחפת מעל כולנו – האם המכונה אשמה, או מי שהפעיל אותה?

כולנו משתמשים ב-AI, רק לא כולנו מבינים איך

האמת היא ש Deloitte לא לבד. [כמעט כל משרד ייעוץ](#), חברת מדיה או גוף ממשלתי משתמש היום בבינה מלאכותית כדי לחסוך זמן, לשפר תוצרים ולהיראות חכמים יותר. זו כבר לא טכנולוגיה ניסיונית, זו השגרה החדשה של העבודה. וזה בסדר. אבל השאלה היא איך עושים שימוש ב-AI-כאן בדיוק מתחילה הבעיה. כשהמכונה שנועדה לסייע בכתיבה או ניתוח נתונים הופכת לחלק מתהליך אנושי שאיבד את היכולת להבחין איפה נגמר האדם ומתחיל האלגוריתם – נוצר ואקום מסוכן.

ב Deloitte-אף אחד לא היה "רע". פשוט לא היה מִישוּ אחראי. לא היה מי שיבדוק, מי שיאשר, מי שישאל את השאלה הבסיסית: האם מה שכתוב כאן באמת נכון? וזה כל ההבדל בין תקלה טכנולוגית לתקלה ארגונית. כשאנחנו מפקידים בידי AI משימה מקצועית בלי פיקוח ובקרה, אנחנו כבר לא משתמשים בו ככלי, אנחנו משתמשים בו כתחליף לחשיבה.

מה AI עושה כשהוא לא יודע – ומה זה מלמד עלינו

בינה מלאכותית לא משקרת. [היא פשוט ממציאה תשובה כשהיא לא יודעת](#) אם לא אמרנו לה לעצור. אבל בני אדם כן יודעים לעצור. הם יודעים לבדוק, להצליב, לוודא. רק שבמקרה של Deloitte הם לא עשו את זה. וזה אולי השיעור החשוב ביותר בפרשה כולה:

הכשל לא טמון במודל, אלא באחריות האנושית בעולם שבו הגבולות בין אדם למכונה הולכים ומיטשטשים.

AI הוא לא קסם. הוא מראה. הוא לא הופך ארגונים חכמים יותר, הוא חושף עד כמה הם חכמים באמת. ארגון שאין בו נהלים, בקרה והכשרה, יפיק תוצרים שנראים מושלמים מבחוץ אך ריקים מתוכן מבפנים. הוא יראה חכם יותר, כשהוא בעצם רק טועה מהר יותר.

► Auditor admits AI used in \$440k report

Deloitte to pay for AI mistakes

Edmund Tadros and Paul Karp

Deloitte Australia will issue a partial refund to the federal government after admitting that artificial intelligence had been used in the creation of a \$440,000 report littered with errors including three non-existent academic references and a made-up quote from a Federal Court judgement.

A new version of the report for the Department of Workplace Relations (DEWR) was quietly uploaded to the department's website on Friday, ahead of a long weekend across much of Australia. It features more than a dozen deletions of non-existent references and footnotes, a rewritten reference list, and corrections to multiple typographic errors.

The first version of the report, about the IT system used to automate penalties in the welfare system such as pauses on the dole, was published in July. Less than a month later, Deloitte was forced to investigate the report after University of Sydney academic

Tink calls for food waste tax

Kylea Tink, the incoming chief of Australia's largest hunger relief charity, Foodbank, will push the Albanese government to introduce a national food waste tax incentive. **News p2**

Treasury may allow small lenders to exit the bank data sharing regime known as open banking. **Companies p19**

Dr Christopher Rudge highlighted multiple errors in the document.

At the time, Rudge speculated that the errors may have been caused by what is known as "hallucinations" by generative AI. This is where the technology responds to user queries by inventing references and quotes. Deloitte declined to comment.

The incident is embarrassing for Deloitte as it earns a growing part of its \$US70.5 billion (\$107 billion) in annual global revenue by providing advice and training clients and executives about **Continued p8**

6. Background & Methodology

Phase 3: Evaluation and Preliminary Findings

• Traceability and Documentation Gaps – Assessed whether system code state can be mapped to business requirements and compliance needs, which included the use of a generative AI large language model (Azure OpenAI GPT-4o) based tool chain licensed by DEWR and hosted on DEWR's Azure tenancy.

New text added revealing the use of AI to produce the report.

"Deloitte has now issued a confession, albeit buried in the methodology section."

Dr Christopher Rudge, the academic who first uncovered the errors.



Amanda Rishworth's department will receive a refund from Deloitte.

Source: Australian Financial Review

המבחן האמיתי של העידן החדש

המהפכה הגרטיבית לא נמדדת בכמה מהר אנחנו משתמשים בבינה מלאכותית, אלא בכמה חכם, אחראי וזהיר אנחנו עושים את זה. הארגונים שידעו לשלב בין דיוק טכנולוגי לבין אנושיות זהירה, ימצאו את הנוסחה החדשה של אמון. אלה שלא, יגלו שהמכונה לא תחליף אותם – היא פשוט תחשוף כמה קל היה להחליף אותם מלכתחילה.

כי בסוף, לא המכונה נכשלה, אלא האנשים ששכחו לבדוק את מה שהיא יצרה לפני ששלחו את זה הלאה.